



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/695,979	10/29/2003	Raimo Bakis	YOR920030379US01	8009
48175	7590	03/04/2009		
BMT/IBM 50 LOCUST AVENUE NEW CANAAN, CT 06840			EXAMINER GODBOLD, DOUGLAS	
			ART UNIT	PAPER NUMBER
			2626	
			MAIL DATE	DELIVERY MODE
			03/04/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/695,979	Applicant(s) BAKIS ET AL.	
	Examiner DOUGLAS C. GODBOLD	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 04 December 2008.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-37 and 41-44 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-37 and 41-44 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. This Office Action is in response to correspondence filed December 4, 2008 in reference to application 10/695,979. Claims 1-37 and 41-44 are pending and have been examined.

Response to Amendment

2. The amendment filed December 4, 2008 has been accepted and considered in this office action. Claim 17 has been amended.

Response to Arguments

3. Applicant's arguments filed December 4, 2008 have been fully considered but they are not persuasive.

4. With regards to applicants arguments, see remarks pages 10 and 11, regarding the rejections under 35 U.S.C. 112, the examiner disagrees. For instance claim 15 recites "certain amount" however there is no indication as to what exactly a "certain amount" is. Again in claim 16, no definition is given for what qualifies as "expertise." Therefore the claims are indefinite and were properly rejected under section 112.

5. With regards to claim 1, the applicant argues that Henton does not teach selecting style sheets, see arguments page 12. As noted in the previous rejection,

Art Unit: 2626

Henton was not relied upon to teach style sheets, only Nielson was used to teach these limitations.

Further regarding claim 1, applicant argues, see remarks pages 12-15 that Nielson fails to teach the limitations of "said speech style sheet defining desired speech characteristics for a first voice style associated with a first voice-type, said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type, speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type," the examiner disagrees. The applicant specifically argues that it would not have been obvious to associate "Susan" and "woman" with H1. The examiner made this assertion to point out that it is not beyond the scope of the invention to have two voice types for another heading such as H1. It would have been well within the scope of Nielsen to have the voice families "Fred" and "Man" associated with H1. The styles sheet in figure 5 is just an example, not the only embodiment of Nielson. Column 6 lines 25-32 describe listing two voice types, the second as a backup. It is not outside the scope of Nielson to apply this to another section such as H1. Even assuming, arguendo, that having two voice properties associated with H1 would be outside the scope of what is taught by Nielson, it would have been an exceedingly obvious modification, obviated by the motivation given in column 6 line 30, that this step would ensure the speech to be properly universally by all synthesizers.

Art Unit: 2626

6. Applicant's arguments regarding claim 17, see Remarks pages 15-18 are similar in content to that of claim 1. The examiner disagrees for the same reasons given in the response regarding claim 1.

7. Applicant's arguments regarding claim 24, see Remarks pages 18-20 are similar in content to that of claim 1. The examiner disagrees for the same reasons given in the response regarding claim 1.

8. Applicant's arguments regarding claim 29, see Remarks pages 20-22 are similar in content to that of claim 1. The examiner disagrees for the same reasons given in the response regarding claim 1.

9. Applicant's arguments regarding claim 32, see Remarks pages 22-25 are similar in content to that of claim 1. The examiner disagrees for the same reasons given in the response regarding claim 1.

10. Applicant's arguments regarding claim 33, see Remarks pages 26-28 are similar in content to that of claim 1. The examiner disagrees for the same reasons given in the response regarding claim 1.

Claim Rejections - 35 USC § 101

11. 35 U.S.C. 101 reads as follows:

Art Unit: 2626

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

12. Claims 18-23 are rejected under 35 U.S.C. 101 because the claimed invention lacks patentable utility. Claims 18-23 are directed towards a "speech style sheet."

There is no input or output, nor is there a transformation of any kind, and therefore the claimed subject matter has no utility on its own. Therefore claims 17-23 are rejected as lacking utility.

13. Claim 23 is rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. Claim 23 further defines the speech style sheet as one of a programming object, a programming module, a computer program, or a computer files, all of which are computer code. Computer code does not fall into any of the statutory categories. Therefore claims 23 is further rejected as being non-statutory subject matter under 35 U.S.C 101.

Claim Rejections - 35 USC § 112

14. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

15. Claims 15, 16, 17, 37 is rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Art Unit: 2626

16. The term "less" in claim 15 is a relative term which renders the claim indefinite. The term "less" is not defined by the claim, the specification does not provide a standard for ascertaining the requisite degree, and one of ordinary skill in the art would not be reasonably apprised of the scope of the invention.

17. The term "not having expertise in voice arts" in claim 16 is a relative term which renders the claim indefinite. The term "not having expertise in voice arts" is not defined by the claim, the specification does not provide a standard for ascertaining the requisite degree, and one of ordinary skill in the art would not be reasonably apprised of the scope of the invention.

18. Claim 17 recites the limitation "the second voice type" in line 7 of the claim. There is insufficient antecedent basis for this limitation in the claim. Therefore claims 17 is rejected as being indefinite.

19. Claim 37 recites the limitations "said particular gender, said language, and said accent, and said another accent.". There is insufficient antecedent basis for this limitation in the claim.

Claim Rejections - 35 USC § 103

20. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

21. Claims 1-33, 36, and 43 are rejected under 35 U.S.C. 103(a) as being unpatentable Henton (US Patent 5,860,064) over in view of Nielsen (US Patent 5,899,975).

22. Consider claim 1, Henton teaches a method (figure 5), comprising:
identifying text to convert to speech (select text, step 501);
marking said text to associate said text with a desired speech style (figures 2-4 show marking text with colors, size, and boldface in order to associate text with a speech style); and
converting said text to speech having said desired speech characteristics by applying a low level markup generated by a speech style sheet (Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507.).

But Henton does not specifically teach selecting a speech style sheet from a set of available speech style sheets, said speech style sheet defining desired speech characteristics for a first voice style associated with a first voice-type, said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type, speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type;

Art Unit: 2626

In the same field of Speech presentation, Nielsen teaches selecting a speech style sheet from a set of available speech style sheets (style sheets are selected, based on author specified, or local user; column 7 lines 1-23), said speech style sheet (figure 5) defining desired speech characteristics for a first voice style associated with a first voice-type (figure 5, voice type is defined by the "Body" class and associated variables being set, and voice style is Susan.), said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type (second voice style is woman, which serves as a backup to Susan; column 6 lines 24-33), speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type (Figure 5, H1 could be the 2nd voice type. Although not specifically shown, it is obvious that "Susan" and "woman" could be specified for H2 as well. It is described column 6 lines 24-33, when a voice family is specified, two may be listed, one as a backup. Given this, and the flexibility of the Speech Style sheet, it is obvious that two types could in fact contain the same two styles.)

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speech style sheets of Nielsen to specify the parameters for the speech synthesis of Henton in order to allow customization of reading style in a way that can be easily transmitted over networks and can be easily be used in web based applications or used on different speech synthesizers.

Art Unit: 2626

23. Consider claim 2, Henton teaches a method according to claim 1, further comprising:

sending said text with said low level markup to an output device (Obtained vocal parameters will be outputted by the text to speech system; column 4, line 45. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42.).

24. Consider claim 3, Henton and Nielsen teaches a method according to claim 1, further comprising:

identifying at least one low level markup (Henton, columns of Table 2);

defining a voice style at least in part by associating said voice style with said at least one low level markup (Henton, Table 2 gives examples of the defined emotions of the preferred embodiment of the present invention with their associated vocal emotion values; column 9, line 56.); and

associating a speech style sheet with said voice style (Nielsen, Figures 5 and 6 are examples of speech style sheets that specify a voice style; Columns 6 and 7.).

25. Consider claim 4, Nielsen teaches a method according to claim 3, wherein said associating said speech style sheet with said voice style includes:

creating said speech style sheet (Figure 6 is a style sheet generated by a user; column 6, line 48).

Art Unit: 2626

26. Consider claim 5, Nielsen and Henton teaches a method according to claim 3, wherein said associating said speech style sheet with said voice style includes:

editing said speech style sheet (Nielsen: Figure 6 is a style sheet generated by a user; column 6, line 48. Henton: As such, note that the particular values shown are easily modifiable, by the system implementer and/or the user, to thus allow for differences in cultural interpretations and user/listener perceptions; column 9, line 61.).

27. Consider claim 6, Henton and Nielsen teaches a method according to claim 1, wherein said low level markup defines at least one of a pitch, a prosody, a voice quality, a duration, a tremor, a timbre, a speed, an intonation, a timing, a volume, and a pronunciation rule (Table 2 gives examples of the defined emotions of the preferred embodiment of the present invention with their associated vocal emotion values; column 9, line 56. Table 2, shows pitch mean, range, volume, and speaking rate; Nielsen also shows some of these values in the style sheets of figures 5 and 6.).

28. Consider claim 7, Henton and Nielsen teaches a method according to claim 1, further comprising:

providing said speech style sheet to at least one of a text-to-speech developer and a text-to-speech device (Henton: As such, note that the particular values shown are easily modifiable, by the system implementer and/or the user, to thus allow for differences in cultural interpretations and user/listener perceptions; column 9, line 61. Style sheets must be presented to a developer to be modified. Obtained vocal

Art Unit: 2626

parameters will be outputted by the text to speech system; column 4, line 45. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42. Nielsen, speech style sheets used to synthesize speech; column 7 lines 31-47.).

29. Consider claim 8, Nielsen teaches a method according to claim 1, further comprising:

compiling a library of speech style sheets. (Style sheet database 240, figure 2; column 4 line 50. This must be compiled to exist.)

30. Consider claim 9, Henton and Nielsen teaches a method according to claim 1, further comprising:

identifying at least one low level markup (column 11 lines 28-35 show text marked up with low level parameters.);

associating a speech style sheet with said at least one low level markup (Henton; Column 11 lines 28-35 show text marked up with low level parameters that were a result of applying different vocal emotions [from table 2] to different portions of text; column 11, line 1. Nielsen; style sheets of figures 5 and 6 show low level markup parameters such as voice-pitch, pitch volume, etc.).

31. Consider claim 10, Nielsen teaches a method according to claim 1, wherein said speech style sheet is selected from a menu of available speech style sheets (style

Art Unit: 2626

sheets can be specified from a user; column 6 line 15. presumably, out of a database such as Figure 2, database 240.).

32. Consider claim 11, Henton teaches a method according to claim 1, wherein said marking of said text includes annotating said text with an annotation such as underlining, bolding, italicizing, highlighting, color-coding, coding, adding a symbol, a mark, or a design (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; column 9, line 7.).

33. Consider claim 12, Henton and Nielsen teaches a method according to claim 1, wherein said converting said text to speech includes:

identifying said low level markup associated with said speech style sheet
(Column 11 lines 28-35 show text marked up with low level parameters that were a result of applying different vocal emotions [from table 2] to different portions of text; column 11, line 1.); and

converting said marking of said text to said low level markup (Figures 2-4, text is marked using color codes to determine an emotion; described in detail column 7 line 60- column 9 line 11. Figure 5, Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507. Final marked up text with emotion values shown in column 11, line 28-35. When combined with Nielsen, it would be obvious that the low level can be pulled from the style sheet, instead of a table.).

34. Consider claim 13, Henton and Nielsen teaches a method according to claim 1, wherein said marking of said text further associates said text with a voice style associated with said speech style sheet (Figures 2-4, text is marked using color codes to determine an emotion; described in detail column 7 line 60-column 9 line 11. Emotions and parameters are shown in table 2. When combined with Nielsen, it would be obvious that the low level can be pulled from the style sheet, instead of a table.).

35. Consider claim 14, Henton and Nielsen teaches a method according to claim 13, wherein said voice style represents at least one of an age, an educational level, an emotion, a feeling, a physical trait, a personality trait, and a speech category (Henton teaches a method for automatic application of vocal emotion parameters, abstract. Nielsen shows man and woman in figures 5 and 6.).

36. Consider claim 15, Henton teaches a method according to claim 1, wherein said low level markup allows a text-to-speech developer to convey a certain amount of information using less text. (Column 11 lines 28-35 show text marked up with low level parameters that were a result of applying different vocal emotions [from table 2] to different portions of text; column 11, line 1. These low level parameters convey information using text to the synthesizer.).

Art Unit: 2626

37. Consider claim 16, Henton teaches a method according to claim 1, wherein said selecting is performed by a text-to-speech developer not having expertise in voice arts (What is needed, therefore, is an intuitive graphical interface for specification and modification of vocal emotion of synthetic speech; column 2, line 36. Further, the present invention provides for the automatic specification of prosodic controls which create vocal emotional affect in synthetic speech produced with a concatenative speech synthesizer, column 2, line 64.).

38. Consider claim 17, Nielsen teaches An article of manufacture comprising:

39. a computer usable medium having computer readable code embodied (figure 2) therein, the computer readable code defining a speech style sheet (Figures 5 and 6), comprising:

speech characteristics for at least one voice style associated with at least one voice-type (figure 5, voice type is defined by the "Body" class and associated variables being set, and voice style is Susan), said speech characteristics for at least one voice style associated with said at least one voice-type including: speech characteristics for a first voice style associated with a first voice-type (figure 5, voice type is defined by the "Body" class and associated variables being set, and voice style is Susan), speech characteristics for a second voice style associated with the first voice-type (second voice style is woman, which serves as a backup to Susan; column 6 lines 24-33), speech characteristics for the first voice style associated with the second voice-type, and speech characteristics for the second voice style associated with the second voice-

Art Unit: 2626

type (Figure 5, H1 could be the 2nd voice type. Although not specifically shown, it is obvious that "Susan" and "woman" could be specified for H2 as well. It is described column 6 lines 24-33, when a voice family is specified, two may be listed, one as a backup. Given this, and the flexibility of the Speech Style sheet, it is obvious that two types could in fact contain the same two styles.).

Nielson does not specifically teach said at least one voice style relating a high level markup of said voice-type to a low level markup of said voice-type.

In the same field of speech to text, Henton teaches at least one voice style relating a high level markup of said voice-type to a low level markup of said voice-type (Device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. Associations are shown in table 2. Figures 2-4 show marking up text using color coding, bolding, and font size to associate emotions with text for emotions; column 9, line 7.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the high level to low level conversion of Henton with the style sheets of Neilson in order to allow for intuitive manipulation by a user; Henton column 2 line 56.

40. Consider claim 18, Henton teaches the speech style sheet according to claim 17, wherein said high level markup of said voice-type is a text markup (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

41. Consider claim 19, Henton teaches the speech style sheet according to claim 17, wherein said high level markup includes at least one of an underlining, a bolding, an italicizing, a highlighting, a color-coding, an annotation, a coding, and an application of at least one of a symbol, a mark, and a design (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

42. Consider claim 20, Henton and Nielsen teaches the speech style sheet according to claim 17, wherein said low level markup of said voice-type includes code causing generation of speech having particular speech properties (Henton, Column 11 lines 28-35 show text marked up with low level parameters that were a result of applying different vocal emotions [from table 2] to different portions of text; column 11, line 1. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42. When combined with Nielsen, it would be obvious that the low level can be pulled from the style sheet, instead of a table.).

43. Consider claim 21, Henton and Nielsen teaches the speech style sheet according to claim 17, wherein said low level markup defines at least one of a pitch, a prosody, a voice quality, a duration, a tremor, a timbre, speed, an intonation, a timing, a volume, and a pronunciation rule (Henton, Table 2 gives examples of the defined emotions of the preferred embodiment of the present invention with their associated vocal emotion

Art Unit: 2626

values; column 9, line 56. Table 2, shows pitch mean, range, volume, and speaking rate. Nielsen; style sheets of figures 5 and 6 show low level markup parameters such as voice-pitch, pitch volume, etc).

44. Consider claim 22, Henton and Nielsen teaches the speech style sheet according to claim 17, wherein said at least one voice style represents style characteristics such as an age, an educational level, an emotion, a feeling, a physical trait, a personality trait, and a speech category (Henton teaches a method for automatic application of vocal emotion parameters, abstract. Nielsen shows man and woman in figures 5 and 6.).

45. Consider claim 23, Nielsen teaches the speech style sheet according to claim 17, wherein said speech style sheet is at least one of a programming object, a programming module, a computer program, or a computer file (Figure 2, database 204 stores style sheets. therefore they must be stored as a computer file.).

46. Consider claim 24, Henton teaches an apparatus (figure 1), comprising:
a processor having access to at least one speech style definition, and said definition relating a high level markup of said voice-type to a low level markup of said voice-type (Device contains a memory for holding said vocal emotions parameters associated with emotions, column 4, line 54. Associations are shown in table 2. Figures 2-4 show marking up text using color coding, bolding, and font size to associate

Art Unit: 2626

emotions with text for emotions; column 9, line 7.), wherein said processor is operative to convert said high level markup to said low level markup (Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507.);

a user interface device for applying said at least one voice style to text associated with said voice-type, said user interface being in communication with said processor (Figure 1, a keyboard 13, or other textual input device such as a write-on tablet or touch screen, provides input to the CPU/memory unit 11, as does input controller 15 which by way of example can be a mouse, a 2-D trackball, a joystick, etc.; column 5, line 22.); and

an output device connected to said processor for converting said text with said low level markup to speech (figure 1, output 21. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42.).

But Henton does not specifically teach a processor having access to at least one speech style sheet, said at least one speech style sheet containing a definition of a voice style associated with a voice-type said speech style sheet defining desired speech characteristics for a first voice style associated with a first voice-type, said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type, speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type.

Art Unit: 2626

In the same field of Speech Synthesizers, Nielsen teaches a processor having access to at least one speech style sheet (style sheets are selected, based on author specified, or local user; column 7 lines 1-23); said at least one speech style sheet containing a definition of a voice style associated with a voice-type said speech style sheet defining desired speech characteristics for a first voice style associated with a first voice-type (figure 5, voice type is defined by the "Body" class and associated variables being set, and voice style is Susan), said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type (second voice style is woman, which serves as a backup to Susan; column 6 lines 24-33), speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type (Figure 5, H1 could be the 2nd voice type. Although not specifically shown, it is obvious that "Susan" and "woman" could be specified for H2 as well It is described column 6 lines 24-33, when a voice family is specified, two may be listed, one as a backup. Given this, and the flexibility of the Speech Style sheet, it is obvious that two types could in fact contain the same two styles).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speech style sheets of Nielsen to specify the parameters for the speech synthesis of Henton in order to allow customization of reading style in a way that can be easily transmitted over networks and can be easily be used in web based applications or used on different speech synthesizers.

Art Unit: 2626

47. Consider claim 25, Henton teaches the apparatus of claim 24, wherein said processor includes at least one of a text-to-speech engine (The preferred manner in which this invention would be implemented is in the context of creating vocal emotions that may be associated with text that is to be read by a text-to-speech synthesizer; column 9, line 15.) and a text normalizer (a simple linear normalization is then performed in the preferred embodiment of the present invention in order to translate the graphical modifications to the resulting vocal emotion effect; column 9, line 38).

48. Consider claim 26, Henton and Nielsen teaches the apparatus according to claim 24, wherein said low level markup defines at least one of a pitch, a prosody, a voice quality, a duration, a tremor, a timbre, a speed, an intonation, a timing, a volume, and a pronunciation rule (Table 2 gives examples of the defined emotions of the preferred embodiment of the present invention with their associated vocal emotion values; column 9, line 56. Table 2, shows pitch mean, range, volume, and speaking rate. Nielsen; style sheets of figures 5 and 6 show low level markup parameters such as voice-pitch, pitch volume, etc).

49. Consider claim 27, Henton teaches the apparatus according to claim 24, wherein said high level markup includes at least one of an underlining, a bolding, an italicizing, a highlighting, a color-coding, an annotation, a coding, and an application of at least one of a symbol, a mark, and a design (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

50. Consider claim 28, Henton and Nielsen teaches the apparatus according to claim 24, wherein said voice style represents at least one of an age, an educational level, an emotion, a feeling, a physical trait, a personality trait, and a speech category (Henton teaches a method for automatic application of vocal emotion parameters, abstract. Nielsen shows man and woman in figures 5 and 6.).

51. Consider claim 29, Henton teaches a system (Figure 1), comprising:
a text-to-speech device for receiving text associated with a voice-type (The preferred manner in which this invention would be implemented is in the context of creating vocal emotions that may be associated with text that is to be read by a text-to-speech synthesizer; column 9, line 15.), said text having a high level markup associated with said voice style (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.), said text-to-speech device having access to said speech style sheet (CPU 11, connected to memory 17. Memory holds vocal emotion parameters associated with emotions; column 4, line 54.) and also having:

a memory for storing computer executable code (figure 1, memory 17);

and

a processor for executing the program code stored in memory (CPU 11), wherein the program code includes;

code to determine, by accessing said speech style sheet, a low level markup associated with said high level markup (Figure 5, Look up synthesizer values for chosen emotion in emotion table [table 2], step 505.); and

code to convert said high level markup of said text to said low level markup (Apply speech synthesizer vocal emotion values to the chosen text, step 507.); and

an output device for producing expressive speech using said text with said low level markup, said output device in communication with said text-to-speech device (figure 1, output 21. Values shown in Table 2 are input to the speech synthesizer, Column 10, line 42.)

Henton does not specifically teach:

a designer device for creating speech style sheets;

a speech style sheet at least partially created by said designer device, said speech style sheet defining desired speech characteristics for a first voice style associated with a first voice-type, said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type, speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type.

In the same field of Speech Synthesizers, Nielsen teaches:

a designer device for creating speech style sheets (Figure 6 is a style sheet generated by a user; column 6, line 48. It is inherent this must be done on a device.)

Art Unit: 2626

a speech style sheet at least partially created by said designer device (Figure 6 is a style sheet generated by a user; column 6, line 48), said speech style sheet defining desired speech characteristics for a first voice style associated with a first voice-type (figure 5, voice type is defined by the "Body" class and associated variables being set, and voice style is Susan), said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type (second voice style is woman, which serves as a backup to Susan; column 6 lines 24-33), speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type (Figure 5, H1 could be the 2nd voice type. Although not specifically shown, it is obvious that "Susan" and "woman" could be specified for H2 as well. It is described column 6 lines 24-33, when a voice family is specified, two may be listed, one as a backup. Given this, and the flexibility of the Speech Style sheet, it is obvious that two types could in fact contain the same two styles).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speech style sheets of Nielsen to specify the parameters for the speech synthesis of Henton in order to allow customization of reading style in a way that can be easily transmitted over networks and can be easily be used in web based applications or used on different speech synthesizers.

52. Consider claim 30, Henton teaches the system according to claim 29, further comprising:

Art Unit: 2626

a developer device in communication with said text-to-speech device (Figure 1, a keyboard 13, or other textual input device such as a write-on tablet or touch screen, provides input to the CPU/memory unit 11, as does input controller 15 which by way of example can be a mouse, a 2-D trackball, a joystick, etc.; column 5, line 22.), said developer device for marking text and providing said text to said text-to-speech device (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

53. Consider claim 31, Henton teaches the system according to claim 29, further comprising:

a user interface device in communication with said text-to-speech device (Figure 1, a keyboard 13, or other textual input device such as a write-on tablet or touch screen, provides input to the CPU/memory unit 11, as does input controller 15 which by way of example can be a mouse, a 2-D trackball, a joystick, etc.; column 5, line 22.), said user interface device for applying high level markup to text and providing said text to said text-to-speech device (Figures 2-4 show marking up text using color coding, bolding, and font size changes for emotions; columns 7 line 61 - 9, line 11.).

54. Consider claim 32, Henton teaches an article of manufacture (figure 1), comprising:

a computer usable medium having computer readable program code means embodied therein for producing expressive text-to-speech (External storage 17, which

Art Unit: 2626

can include fixed disk drives, floppy disk drives, memory cards, etc., is used for mass storage of programs and data; column 5, line 26. Method, figure 5.), comprising:

computer readable program code means for identifying text to convert to speech (select text, step 501);

computer readable program code means for marking said text to associate said text with a desired speech style (figures 2-4 show marking text with colors, size, and boldface in order to associate text with a speech style); and

computer readable program code means for converting said text to speech having said desired speech characteristics by applying a low level markup generated by a speech style sheet (Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507.).

But Henton does not specifically teach computer readable program code means for selecting a speech style sheet from a set of available speech style sheets, said speech style sheet defining desired speech characteristics for a first voice style associated with a first voice-type, said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type, speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type;

In the same field of Speech presentation, Nielsen teaches computer readable program code means for selecting a speech style sheet from a set of available speech style sheets (style sheets are selected, based on author specified, or local user; column

Art Unit: 2626

7 lines 1-23), said speech style sheet (figure 5) defining desired speech characteristics for a first voice style associated with a first voice-type (figure 5, voice type is defined by the "Body" class and associated variables being set, and voice style is Susan.), said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type (second voice style is woman, which serves as a backup to Susan; column 6 lines 24-33), speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type (Figure 5, H1 could be the 2nd voice type. Although not specifically shown, it is obvious that "Susan" and "woman" could be specified for H2 as well. It is described column 6 lines 24-33, when a voice family is specified, two may be listed, one as a backup. Given this, and the flexibility of the Speech Style sheet, it is obvious that two types could in fact contain the same two styles.)

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speech style sheets of Nielsen to specify the parameters for the speech synthesis of Henton in order to allow customization of reading style in a way that can be easily transmitted over networks and can be easily be used in web based applications or used on different speech synthesizers.

55. Consider claim 33, Henton teaches a system for producing expressive text-to-speech, (system figure 1, Method figure 5), comprising:

means for identifying text to convert to speech (select text, step 501);

means for marking said text to associate said text with a desired speech style (figures 2-4 show marking text with colors, size, and boldface in order to associate text with a speech style); and

means for converting said text to speech having said desired speech characteristics by applying a low level markup generated by a speech style sheet (Look up synthesizer values for chosen emotion in emotion table [table 2], step 505. Apply speech synthesizer vocal emotion values to the chosen text, step 507.).

But Henton does not specifically teach means for selecting a speech style sheet from a set of available speech style sheets, said speech style sheet defining desired speech characteristics for a first voice style associated with a first voice-type, said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type, speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type;

In the same field of Speech presentation, Nielsen teaches means for selecting a speech style sheet from a set of available speech style sheets (style sheets are selected, based on author specified, or local user; column 7 lines 1-23), said speech style sheet (figure 5) defining desired speech characteristics for a first voice style associated with a first voice-type (figure 5, voice type is defined by the "Body" class and associated variables being set, and voice style is Susan.), said speech style sheet further defining speech characteristics for a second voice style associated with the first voice-type (second voice style is woman, which serves as a backup to Susan; column 6

Art Unit: 2626

lines 24-33), speech characteristics for the first voice style associated with a second voice-type, and speech characteristics for the second voice style associated with the second voice-type (Figure 5, H1 could be the 2nd voice type. Although not specifically shown, it is obvious that "Susan" and "woman" could be specified for H2 as well. It is described column 6 lines 24-33, when a voice family is specified, two may be listed, one as a backup. Given this, and the flexibility of the Speech Style sheet, it is obvious that two types could in fact contain the same two styles.)

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speech style sheets of Nielsen to specify the parameters for the speech synthesis of Henton in order to allow customization of reading style in a way that can be easily transmitted over networks and can be easily be used in web based applications or used on different speech synthesizers.

56. Consider claim 36, Henton teaches the speech style sheet according to claim 17, wherein said language is English (All examples in figures 204 are in English.)

57. Consider claim 43, Henton and Nielsen teach the method according to claim 1, wherein:

said first voice style represents at least one of an age, an educational level, an emotion, a feeling, a physical trait, a personality trait (Nielsen; figure 5, Woman or Susan given);

Art Unit: 2626

said second voice style represents at least one of an age, an educational level, an emotion, a feeling, a physical trait, a personality trait (Nielsen; figure 5, Woman or Susan given;

said first voice-type represents a voice speaking in a language (all examples in Nielsen and Henton are in English); and

said second voice-type represents a voice speaking in a language (all examples in Nielsen and Henton are in English).

58. Claim 34 is rejected under 35 U.S.C. 103(a) as being unpatentable over Henton in view of Nielsen as applied to claims 1 and 24 above, and further in view of Atkin et al (US PAP 2004/0260551).

59. Consider claim 34, Henton in view of Nielsen teaches the method according to claim 1, but does not specifically teach wherein said selected speech style sheet defines pronunciation rules for at least one of aviation, chemistry and real estate.

However in the same field of speech to text, Atkin suggests said selected speech style sheet defines pronunciation rules for at least one of aviation, chemistry and real estate (A subject matter semantic identifier corresponds to particular subject matter, such as a children's book or a financial article. A user interest semantic identifier corresponds to particular areas of interest, such as a summary, detail, or section headings of a text file. For example, the semantic analyzer identifies that a text block is a paragraph corresponding to financial information and associates a "Business Journal"

Art Unit: 2626

semantic identifier with the text block. In this example, the semantic analyzer retrieves voice attributes corresponding to the "Business Journal" semantic identifier from the look-up table. The semantic analyzer provides the voice attributes to a voice reader. The voice attributes include attributes such as a pitch value, a loudness value, and a pace value. In one embodiment, the voice attributes are provided to the voice reader through an Application Program Interface (API). The voice reader inputs the voice attributes into a voice synthesizer whereby the voice synthesizer converts the text block into synthesized speech for a user to hear; paragraphs 0010 and 0011. Although it does not specifically say aviation or chemistry or real estate, one of ordinary skill in the art could appreciate that this process is applicable to these fields as well.).

Therefore it would have been obvious to one of ordinary skill in the art to use the context dependency as taught by Atkin with the style sheets of Henton in view of Neilsen in order to provide a context dependent speech synthesizer.

60. Claim 35 is rejected under 35 U.S.C. 103(a) as being unpatentable over Henton in view of Nielsen as applied to claim 1 above, and further in view of Surace et al (US Patent 6,334,103.).

61. Consider claim 35, Henton in view of Nielsen teaches the method according to claim 1, but does not teach specifically wherein said selected speech style sheet defines pronunciation rules for an automated flight reservation system.

Art Unit: 2626

In the same field of speech synthesis, Surace suggests said selected speech style sheet defines pronunciation rules for an automated flight reservation system. (In one embodiment, controlling the voice user interface includes providing the voice user interface with multiple personalities. The voice user interface with personality installs a prompt suite for a particular personality from a prompt repository that stores multiple prompt suites, in which the multiple prompt suites are for different personalities of the voice user interface with personality; column 2, line 12. Although this art does not specifically teach a flight reservation, one of ordinary skill in the art can appreciate that a prompting voice system can be used as a flight reservation system.)

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a voice interface with personality as taught by Surace as an application for the style sheet system of Henton in view of Nielsen in order to provide a personalized experience in a voice response system.

62. Claims 37 and 44 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nielsen in view of Henton as applied to claim 17 above, and further in view of Baba (US Patent 6,397,183).

63. Consider claim 37, Henton and Nielsen suggest the speech style sheet according to claim 17, wherein said particular gender is male (Henton, Table 2 values are for a female voice, for a male voice the table values are to be altered, column 10,

Art Unit: 2626

line 1.), said language is common English (Henton, all examples in figures 2-4 are in English).

Henton and Nielsen does not specifically teach: said accent is a southern U.S. accent and said another accent is a Cornish accent

In the same field of text to speech, Baba teaches:

said accent is a southern U.S. accent and said another accent is a Cornish accent (It would be highly desirable to be able to capture a particular style, such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent); column 1, line 28. Although a Cornish accent is not specifically taught, it would be obvious to one of ordinary skill in the art that one could be included in the available styles.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the accents of Baba with the text to speech of Henton and Nielsen in order to allow a user more options when assigned voice types.

64. Consider claim 44, Henton and Nielsen teaches the method according to claim 1, wherein said first voice-type represents a voice of a particular, and wherein said second voice-type represents a voice of said particular gender speaking (Figure 5 shows H1 and Body with men and Women) but does not specifically teach wherein said first voice-type represents a voice of a particular gender speaking in a language with an accent, and wherein said second voice-type represents a voice of said particular gender speaking in said language with another accent.

Art Unit: 2626

In the same field of text to speech, Baba teaches said first voice-type represents a voice of a particular gender speaking in a language with an accent, and wherein said second voice-type represents a voice of said particular gender speaking in said language with another accent (would be highly desirable to be able to capture a particular style, such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent); column 1, line 28. Although a Cornish accent is not specifically taught, it would be obvious to one of ordinary skill in the art that one could be included in the available styles).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the accents of Baba with the text to speech of Henton and Nielsen in order to allow a user more options when assigned voice types.

65. Claims 41 and 42 rejected under 35 U.S.C. 103(a) as being unpatentable over Henton in view of Nielsen as applied to claims 1 and 17 above, and further in view of Kochanski et al. (US Patent 6,810,378).

66. Consider claim 41, Henton and Nielsen teach the method according to claim 1, but does not specifically teach wherein said selected speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category.

In the same field of speech synthesis, Kochanski teaches selected speech style sheet defines pronunciation rules for a speech category and wherein another speech style sheet from said set of available speech style sheets defines pronunciation rules for another speech category (It would be highly desirable to be able to capture a particular style, such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent). This is a pronunciation rule; column 1, line 28. When combined with Baba, it would be obvious to make this a choice for each row in figure 2, or each style sheet.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speaking styles that include accents which include pronunciation information of Kochanski with the style sheets of Henton and Nielsen in order to provide a more robust and flexible speech synthesis device.

67. Consider claim 42, Henton and Nielsen teach the speech style sheet according to claim 1, but does not specifically teach wherein said first voice-type represents a voice of a particular gender speaking in a language with an accent, and wherein said second voice-type represents a voice of said particular gender speaking in said language with another accent..

In the same field of speech synthesis, Kochanski teaches wherein said first voice-type represents a voice of a particular gender speaking in a language with an accent, and wherein said second voice-type represents a voice of said particular gender speaking in said language with another accent. (It would be highly desirable to be able

Art Unit: 2626

to capture a particular style, such as, for example, the style of a specifically identifiable person or of a particular class of people (e.g., a southern accent). This is a pronunciation rule; column 1, line 28. When combined with Baba, it would be obvious to make this a choice for each row in figure 2, or each style sheet. The example of this is in English. Baba also shows gender).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the speaking styles that include accents which include pronunciation information of Kochanski with the style sheets of Henton and Nielsen in order to provide a more robust and flexible speech synthesis device.

Conclusion

68. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Art Unit: 2626

Any inquiry concerning this communication or earlier communications from the examiner should be directed to DOUGLAS C. GODBOLD whose telephone number is (571)270-1451. The examiner can normally be reached on Monday-Thursday 7:00am-4:30pm Friday 7:00am-3:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571) 272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DCG

/Patrick N. Edouard/

Supervisory Patent Examiner, Art Unit 2626